Tuberculosis 95 (2015) S133–S139

Contents lists available at ScienceDirect

Tuberculosis

journal homepage: http://intl.elsevierhealth.com/journals/tube

Ancient mycobacterial lipids: Key reference biomarkers in charting the evolution of tuberculosis

David E. Minnikin [a, *], Oona Y.-C. Lee [a], Houdini H.T. Wu [a], Gurdyal S. Besra [a], Apoorva Bhatt [a], Vijayashankar Nataraj [a], Bruce M. Rothschild [b], Mark Spigelman [c], Helen D. Donoghue [d]

[a] Institute of Microbiology and Infection, School of Biosciences, University of Birmingham, Edgbaston, Birmingham, UK
[b] Biodiversity Institute and Departments of Anthropology and Geology, University of Kansas, Lawrence, KS 66045, USA
[c] Kuvin Center for the Study of Infectious and Tropical Diseases and Ancient DNA, Hadassah Medical School, Hebrew University, Jerusalem, Israel
[d] Centres for Clinical Microbiology and the History of Medicine, University College London, London, UK

Keywords:
Tuberculosis
Evolution
Lipids
Biomarkers
Zoonosis

S U M M A R Y

Mycobacterium tuberculosis has a cell envelope incorporating a peptidoglycan-linked arabinogalactan esterified by long-chain mycolic acids. A range of "free" lipids are associated with the "bound" mycolic acids, producing an effective envelope outer membrane. The distribution of these lipids is discontinuous among mycobacteria and such lipids have proven potential for biomarker use in tracing the evolution of tuberculosis. A plausible evolutionary scenario involves progression from an environmental organism, such as *Mycobacterium kansasii*, through intermediate "smooth" tubercle bacilli, labelled *"Mycobacterium canettii"*; cell envelope lipid composition possibly correlates with such a progression. *M. kansasii* and *"M. canettii"* have characteristic lipooligosaccharides, associated with motility and biofilms, and glycosyl phenolphthiocerol dimycocerosates ("phenolic glycolipids"). Both these lipid classes are absent in modern *M. tuberculosis sensu stricto*, though simplified phenolic glycolipids remain in certain current biotypes. Dimycocerosates of the phthiocerol family are restricted to smaller phthiodiolone diesters in *M. kansasii*. Diacyl and pentaacyl trehaloses are present in *"M. canettii"* and *M. tuberculosis*, accompanied in the latter by related sulfated acyl trehaloses. In comparison with environmental mycobacteria, subtle modifications in mycolic acid structures in *"M. canettii"* and *M. tuberculosis* are notable. The probability of essential tuberculosis evolution taking place in Pleistocene megafauna, rather than *Homo sapiens*, is reemphasised.

1. Introduction

Tuberculosis is an ancient disease, whose pre-Holocene history is shrouded in mystery. Analysis of skeletal material has provided evidence for tuberculosis in *Homo sapiens* at up to 9000 years (9 ka) before present (BP) [1], almost back to the start of the Holocene. Travelling back from the Holocene into the cyclical glacial times of the Pleistocene, human skeletal material becomes scarce and no direct evidence for any tuberculosis in *H. sapiens* has been demonstrated. In contrast, distinctive tuberculosis lesions have been recorded in a range of megafauna and other animals from that epoch. Typically, the lesions take the form of undermined articular surfaces, as exemplified by a metacarpal from *Bison antiquus* recovered from Natural Trap Cave, Wyoming [2]. In addition to the bison metacarpal, 19% of 1002 125 ka to 8 ka BP bovid specimens [3] and 52% of 113 38 ka to 10 ka BP mastodon bones [4] had similar lesions indicative of tuberculosis. Bone lesions cannot be considered as complete proof of tuberculosis diagnosis, but the dearth of comparable lesions in bones from *H. sapiens*, over the same time period, is very striking. To resolve this conundrum it has been proposed [2] that *Mycobacterium tuberculosis* may have been principally an animal disease during its early evolution, with transmission to humans occurring later.

* Corresponding author. School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. Tel.: +44 (0)1214158126.
E-mail addresses: d.e.minnikin@bham.ac.uk (D.E. Minnikin), o.y.lee@bham.ac.uk (O.Y.-C. Lee), h.wu.2@bham.ac.uk (H.H.T. Wu), g.besra@bham.ac.uk (G.S. Besra), a.bhatt@bham.ac.uk (A. Bhatt), VXN131@bham.ac.uk (V. Nataraj), bmr@ku.edu (B.M. Rothschild), spigelman@btinternet.com (M. Spigelman), h.donoghue@ucl.ac.uk (H.D. Donoghue).

http://dx.doi.org/10.1016/j.tube.2015.02.009
1472-9792/

The use of amplified DNA sequences to diagnose tuberculosis in archaeological material has been developed during the past two decades [5]. Major advances in determining full genomic data have been recently provided by the application of so-called "Next Generation Sequencing" [6] and the more direct "Metagenomic" approach [7]. Informative genomic data have been obtained for specimens stretching back to 9 ka in *H. sapiens* [1] and 17 ka in extinct *B. antiquus* [2] and these diagnoses have been supported by the use of robust lipid biomarkers [1,2,5]. The most diagnostic lipids have been mycolic, mycocerosic and mycolipenic acids and members of the phthiocerol family [1,2,5]. These, and a range of other lipids, are vital components in the integrity of the cell envelopes of the tubercle bacillus and related taxa [8]. Their distribution, however, is discontinuous and changes in lipid composition and structure may well be important factors in the evolution of effective pathogenic species. The aim of this paper is to outline a rational scenario for the evolution of the current *M. tuberculosis* complex from possible environmental candidates. The speculative focus will be on correlating changes in cell envelope lipid composition with developing pathogenicity, taking into account the suitability of particular animal hosts along the way. Representative structures of the key lipids under consideration are shown in Figures 1 and 2.

2. An environmental opportunist to a perfect pathogen?

The challenge is to chart a pathway from ancestral environmental freely-circulating mycobacterial species to *M. tuberculosis sensu stricto*, an obligate pathogen with no environmental niche. Currently favoured hypotheses all point to an evolutionary bottleneck, initiated around 35 ka BP [9,10]. Subsequent to this time period, the evolution of a range of particular clades follows an almost linear clonar evolutionary pattern, with key deletions leading to the well-defined modern *M. tuberculosis* complex (MTBC) causing tuberculosis in humans and various animals [11–13].

There is increasing evidence that, before reaching the discontinuity of the bottle-neck, extensive horizontal gene transfer (HGT) was taking place in ancestral tuberculosis strains [10,14]. These strains may not necessarily have been obligate pathogens but opportunist mycobacteria with the ability to survive in the hostile environment of an animal stomach. The rich flora of multiple animal stomachs would provide plentiful opportunities for HGTs, eventually resulting in interim organisms with an enhanced potential to cause tuberculosis. Prime candidates for such a role are pre-bottle-neck ancestral strains, sometimes termed *"M. prototuberculosis"*, [9] which are associated with the "smooth" colony-forming "Canetti" variants of *M. tuberculosis* [15,16]. "*Mycobacterium canettii*" smooth strains continue to be encountered in isolated cases of tuberculosis, but they are usually confined to certain locations in the Horn of Africa [15].

A case for *Mycobacterium marinum* as the pivotal environmental source organism has been advanced [17], but several key factors mitigate against such a selection. The stereochemistries of the *M. marinum* PDIM waxes and PGLs are completely different from those produced by *M. tuberculosis* and *Mycobacterium kansasii* [8,18]. In addition, the oxygenated mycolic acids of *M. marinum* are not cyclopropanated, in contrast with those from *M. tuberculosis* and *M. kansasii* [18,19]. The environmental organism that phenotypically resembles *M. tuberculosis* most closely is *M. kansasii* and this relationship has been supported by genomic comparisons [20,21]. Cogent arguments have been advanced to associate the evolution of tubercle bacilli with bacteria similar to *M. kansasii*, including indications of HGTs between these taxa [20,21]. Key genes acquired by HGT include those coding for mycobacterial lipids, transferases and proteins related to adaptation to anaerobic conditions [20,21]. *M. kansasii* still causes pulmonary disease in Silesian and South African miners, the bacterium being contracted from water in showers [21]. In developing a coherent evolutionary route, the pathway from *M. kansasii*, through *"M. canettii"*, to *M. tuberculosis* is a good working hypothesis. Changes in lipid composition are potentially very significant, involving the mycolic acids, phthiocerol dimycocerosates (PDIMs), glycosyl phenolphthiocerol dimycocerosates ("phenolic glycolipids", PGLs) (Figure 1), diacyl trehaloses (DATs), pentaacyl trehaloses (PATs) and sulfated acyl trehalose glycolipids (SGLs) (Figure 2) [8].

3. Does lipid evolution parallel *M. tuberculosis* evolution?

3.1. Mycolic acids

Possibly the most deep-lying fundamental differences between the lipids from *M. kansasii* and members of the *M. tuberculosis* complex (MTBC) loosely including *"M. canettii"*, are subtle changes in mycolic acid structure. Mycolates from MTBC have characteristic 24-carbon chains in 2-position, whereas the mycolates from *M. kansasii* and the majority of mycobacteria have principally 22-carbon side chains [8,19]. In addition, the MTBC α-mycolates show a significant shortening of the size of the chain between carbon-3 and the proximal cyclopropane (17 → 13 carbons) and the lengthening of the terminal chain (18 → 20 carbons) beyond the distal cyclopropane unit, as compared with *M. kansasii* (Figure 1A) [19]. The methoxymycolates and ketomycolates of *"M. canettii"* and *M. tuberculosis* (Figure 1A) conform to the general pattern of these components in related mycobacteria, such as *M. kansasii*, but, significantly, these oxygenated mycolates are slightly larger than any others [19].

The balance of the three main types of mycolates is possibly significant; the ratios of the α-, methoxy- and ketomycolates are, respectively, ~10:5:8 for *M. kansasii*, ~10:6:8 for "*M. canettii*" and ~10:5:5 for *M. tuberculosis* [19]. Having half of the proportions as α-mycolates in *M. tuberculosis* may be quite significant. The major all *cis*-cyclopropyl α-mycolates of all three taxa are similar in size, centred around 80 carbons. Those from *M. kansasii* are restricted to 80 and 82 carbons overall, but there are four detailed structural varieties of each giving a heterogeneous mixture of eight distinct α-mycolates [19]. In contrast, the four α-mycolates from "*M. canettii*" and *M. tuberculosis* are all very uniform, the two major C_{78} and C_{80} components being accompanied by minor C_{82} and C_{84} mycolates [19]. It is particularly notable that the central (14-carbon) and distal (20-carbon) meromycolate chains are invariable in the α-mycolates from "*M. canettii*" and *M. tuberculosis* (Figure 1A) [19]. The complex methoxymycolates from *M. kansasii*, totalling eight *cis*- and four *trans*-components, have a *cis:trans* ratio of ~3:2,whereas both "*M. canettii*" and *M. tuberculosis* have an enhanced *cis:trans* ratio of ~3:1. [19]. Somewhat simplified methoxymycolates, with six *cis*- and two *trans*-components are found in "*M. canettii*", simplifying further to mainly a C_{85} and lesser C_{87} *cis*- and a single *trans*-methoxymycolate in *M. tuberculosis* [19]. The *trans*-ketomycolates predominate over the *cis*-forms in *M. kansasii* (~6:1), "*M. canettii*" (~4:1) and *M. tuberculosis* (~3:2, respectively); the latter two have mainly a C_{87} *trans*-ketomycolate accompanied by six very minor variants but this contrasts with a heterogeneous mix of ten in *M. kansasii* [19].

The essence of the above seemingly complex changes is an apparent simplification and tightening up of mycolate composition. Mycolic acids are "cornerstones" of the mycobacterial outer membrane, providing a covalent hydrophobic inner leaflet, facilitating binding of the range of "free lipids" (Figures 1 and 2) that comprise the outer leaflet [8]. Physical studies indicate that ketomycolates appear to have a prime structural role in adopting